The Validity of Child and Adolescent Depression Diagnoses in the Danish Psychiatric Central Research Register

Frederiksen, Line Hofmann ; Bilenberg, Niels ; Andersen, Lene ; Henriksen, Natasha ; Jørgensen, Jan ; Steinhausen, Hans-Christoph ; Wesselhoeft, Rikke

Abstract: OBJECTIVE This study examined the validity of childhood depression diagnoses in the Danish Psychiatric Central Research Register (DPCRR) and identified predictors of validity. METHODS A nationwide random sample of 500 children (6-17 years) diagnosed with depression between 1996 and 2016 was identified in the DPCRR. Psychiatric hospital records were reviewed and rated using an online checklist. The primary outcome was, whether depressive symptoms and functional impairment documented in hospital records justified a depressive disorder diagnosis based on ICD-10 or DSM-5 diagnostic criteria. Diagnostic validity was calculated as the positive predictive value. Binary logistic regression analysis was used to identify potential predictors of diagnostic validity and these were included in a multiple logistic regression. RESULTS Psychiatric hospital records were available for 393 patients (78.6%). The documentation in the records justified an ICD-10 depressive episode diagnosis in 72.8%, and DSM-5 major depressive disorder in 73.3% of the patients registered with a depression diagnosis. We identified three predictors of diagnostic validity; 1) The validity increased almost linearly from 2000-2016 (OR 1.14, 95% CI 1.07-1.20, p<0.001), 2) antidepressant use was associated with increased diagnostic validity (OR 2.27, 95% CI 1.35-3.82, p=0.002), and 3) emergency department admission predicted low diagnostic validity (OR 0.33, 95% CI 0.12-0.93, p=0.036). CONCLUSION Childhood depression diagnoses registered in the DPCRR show a satisfactory validity according to both ICD-10 and DSM-5 diagnostic criteria. Diagnostic validity increased steadily from 2000-2016 and was positively correlated to antidepressant use. Depression diagnoses assigned in emergency departments had low diagnostic validity.

DOI: https://doi.org/10.1111/acps.13258

MISS LINE HOFMANN FREDERIKSEN (Orcid ID : 0000-0002-3736-417X)

DR RIKKE WESSELHOEFT (Orcid ID : 0000-0001-9700-2739)

Article type : Original Article

Corresponding author email id: linehofmann@hotmail.dk

I. Title

The Validity of Child and Adolescent Depression Diagnoses in the Danish Psychiatric Central Research Register

II. Short running title

THE VALIDITY OF CHILDHOOD DEPRESSION DIAGNOSES

III. Names of authors

Line Hofmann Frederiksen[1,2] Niels Bilenberg[1,3], Lene Andersen[4], Natasha Henriksen[4], Jan Jørgensen[5], Hans-Christoph Steinhausen[1,6,7,8], Rikke Wesselhoeft[9,10]

IV. Affiliations

[1] Research Unit of Child and Adolescent Mental Health, Institute for Clinical Research, University of Southern Denmark

[2] OPEN, Open Patient data Explorative Network, Odense University Hospital, Odense, Denmark

[3] Department of Child and Adolescent Mental Health Odense, Mental Health Services in the Region of Southern Denmark

[4] Department of Child and Adolescent Mental Health, Mental Health Services in the Region of Southern Denmark

[5] Child and Adolescent Psychiatry, Private Practice, Odense, Denmark

[6] Child and Adolescent Mental Health Centre, Capital Region Psychiatry, Copenhagen, Denmark

[7] Department of Child and Adolescent Psychiatry, Psychiatric University Hospital, Zurich, Switzerland

[8] Clinical Psychology and Epidemiology, Institute of Psychology, University of Basel, Switzerland

[9] Department of Child and Adolescent Mental Health Aabenraa, Mental Health Services in the Region of Southern Denmark

[10] Clinical Pharmacology and Pharmacy, Department of Public Health, University of Southern Denmark, Denmark

V. Acknowledgements

We would like to thank the child and adolescent psychiatric centres in Denmark for providing access to their archives.

The project was funded by the Lundbeck Foundation and the Independent Research Fund Denmark. The funds had no influence on any part of the planning or execution of the study.

VI. Abstract

Objective: This study examined the validity of childhood depression diagnoses in the Danish Psychiatric Central Research Register (DPCRR) and identified predictors of validity.

Methods: A nationwide random sample of 500 children (6-17 years) diagnosed with depression between 1996 and 2016 was identified in the DPCRR. Psychiatric hospital records were reviewed and rated using an online checklist. The primary outcome was, whether depressive symptoms and functional impairment documented in hospital records justified a depressive disorder diagnosis based on ICD-10 or DSM-5 diagnostic criteria. Diagnostic validity was calculated as the positive predictive value. Binary logistic regression analysis was used to identify potential predictors of diagnostic validity and these were included in a multiple logistic regression.

Results: Psychiatric hospital records were available for 393 patients (78.6%). The documentation in the records justified an ICD-10 depressive episode diagnosis in 72.8%, and DSM-5 major depressive disorder in 73.3% of the patients registered with a depression diagnosis.
We identified three predictors of diagnostic validity; 1) The validity increased almost linearly from 2000-2016 (OR 1.14, 95% CI 1.07-1.20, $p<0.001$), 2) antidepressant use was associated with increased diagnostic validity (OR 2.27, 95% CI 1.35-3.82, $p=0.002$), and 3) emergency department admission predicted low diagnostic validity (OR 0.33, 95% CI 0.12-0.93, $p=0.036$).

Conclusion: Childhood depression diagnoses registered in the DPCRR show a satisfactory validity according to both ICD-10 and DSM-5 diagnostic criteria. Diagnostic validity increased steadily from 2000-2016 and was positively correlated to antidepressant use. Depression diagnoses assigned in emergency departments had low diagnostic validity.

Keywords

Child and adolescent psychiatry, depression, register, validity, diagnoses, epidemiology.

VII. Significant outcomes

- The depression diagnoses registered for children and adolescents in the Danish the Danish Psychiatric Central Research Register was confirmed in three out of four according to ICD-10 and DSM-5 diagnostic criteria.
- There was a time trend towards increasing validity of depression diagnoses from the year 2000 and onwards.

- Diagnostic validity was higher for patients treated with antidepressants and lower for patients diagnosed in the emergency department.

Limitations

- Raters reviewing the psychiatric hospital records were not blinded to patient diagnoses.
- Reviews depended on observations documented in the hospital records.
- A sizeable number of the randomly identified hospital records (21.4%) were not available for the study.

VIII. Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

IX. Main Text

Introduction

Depression is a severe and impairing disorder and one of the leading causes of illness and disability in adolescents globally (1). Still, the knowledge of symptom presentation, comorbidity, and trajectories of depression with an onset in childhood and adolescence is limited compared to that of adult onset depression. Furthermore, the number of children and adolescents diagnosed with clinical depression are increasing in some countries, including Denmark (2, 3).

The Nordic countries have extraordinary opportunities to fill this gap of knowledge due to nation-wide registers (4-8) that provide detailed information on a range of variables like assigned clinical psychiatric disorder diagnoses over time. A study is however not stronger than its weakest link, and therefore it is essential that research information collected from population-based registers has a satisfactory validity.

The Danish Psychiatric Central Research Register (DPCRR) has contributed to the characterization of young Danish individuals with depression (9, 10) and psychiatric disorders in general (2, 9, 11). The DPCRR-was incorporated into the Danish National Patient Register in 1995

(5) and it holds data from all in-patient, out-patient and emergency department contacts to psychiatric hospitals (7, 12). The data includes the type, time, and place of a contact as well as the disorder diagnosis, which is registered according to the International Classification of Diseases 10th revision (ICD-10) (13).

Even though the value of national patient registers for research purposes is indisputable, few studies have examined the validity of psychiatric disorder diagnoses registered in the DPCRR or similar registers in other countries (14-17). Furthermore, the existing studies are characterized by marked variation in methodology (18-21). Most commonly, hospital records from the contact that led to assignment of the specific diagnosis have been reviewed to determine the extent to which the documented observations justified the assigned diagnosis. This approach was used in studies of child and adolescent psychiatric disorders in the DPCRR relating to childhood autism (22), early onset obsessive-compulsive disorder (19), early onset schizophrenia (23), and hyperkinetic disorder (20). No studies have yet examined the validity of depression diagnoses available in a national patient register that were assigned to children and adolescents. This is of uttermost importance, because we need to know more about childhood depression and the population-based registers provide remarkable opportunities for this purpose.

The aims of the study

The primary aim was to determine the validity of depression diagnoses assigned to Danish children (6-17 years) in the Danish Psychiatric Central Research Register, based on ICD-10 criteria for depressive episode and DSM-5 diagnostic criteria for major depressive disorder.

Furthermore, the study aimed to examine whether validity varied relative to sex, age, geography, year of diagnosis, disease severity, in-patient/out-patient status, primary or secondary diagnosis, use of psychotropic medication, or use of validated assessment tools in the diagnostic assessment.

Materials and methods

Sample selection

All Danish children and adolescents aged 6-17 years who received a depressive episode diagnosis during the period from 01.01.1996 until 31.12.2016 were identified in the DPCRR. The depression

diagnoses (ICD-10) included were mild depressive episode (F32.0), moderate depressive episode (F32.1), severe depressive episode without psychotic symptoms (F32.2) and severe depressive episode with psychotic symptoms (F32.3). The study included cases irrespective of whether the depressive episode was the main indication for the psychiatric contact (primary diagnosis) or not the main indication (secondary diagnosis), as well as diagnoses assigned for in-patients, out-patients and emergency department patients.

A total of 9,371 patients from the DPCRR were eligible for inclusion in the study. The *a priori* power analysis (see Statistics) determined that we needed 387 patients for testing the validity of diagnoses, but due to the possibility that some records might be unavailable, a sample size of 500 patients was chosen.

The random sample (N=500) included cases from all five regions of Denmark and the randomisation was conducted by the Danish National Health Data Agency. Patients were excluded if their hospital records were missing, incomplete or unavailable.

Procedure

The Danish Psychiatric Central Research Register provided demographic data (see Table 1) as well as the Civil Person Registration number (CPR number) of each patient included in the study. The CPR number is a unique Danish identification number assigned to all individuals at birth (6) and used as an identifier in all Danish hospital records. The CPR number was used to extract the psychiatric hospital record from the contact with the child and adolescent psychiatric department where the depression diagnosis had been assigned.

Four child and adolescent psychiatrists and the first author (MD) reviewed and rated the hospital records. For this purpose, an online checklist (Appendix A) was developed for data extraction using the RedCap electronic data capture tool (24). The checklist ensured a systematic extraction of all relevant hospital record information regarding depressive symptoms, duration of symptoms, functional impairment, medical treatment, and diagnostic assessment. Any uncertainties regarding specific hospital records were clarified by consensus within the study group.

Prior to initiation of the study, two inter-rater reliability tests were conducted on a random subsample (N=20) of the total study sample. The test sample was not different from the study sample at a statistically significant level in terms of sex (p=0.63), age (p=0.07), year of diagnosis (p=0.063), Depressive episode severity (p=0.68), admission status (p=0.34), and primary vs.

secondary diagnosis (p=0.678). The first inter-rater reliability test showed an almost perfect overall agreement (κ=0.89) in classifying patients as having a depressive episode or not, while the agreement on the specific depressive episode severity sub-type was moderate (κ=0.48). The divergences were discussed at a consensus meeting, ensuring concordance with diagnostic criteria and a satisfactory inter-rater reliability. The second test showed a perfect (κ=1.00) overall agreement in classifying patients as having a depressive episode or not, and the agreement on the specific depressive episode severity sub-type was almost perfect (κ=0.92). All kappa values were better than chance agreement ($p<0.001$).

Statistics

An a priori power analysis determined the number of cases (N) needed to show any validity from 0-100 % (p) with a 95% confidence interval (Z) and a 5 % margin of error (E), using the following formula; $N = (Z^2 * p * (1-p))/E^2$. Inter-rater reliability was analysed using Fleiss's kappa (25), and Landis and Koch's scale was used to evaluate the results (26). Baseline data and data on record availability was analysed using X^2 or Fisher's exact tests for categorical variables and Kruskal Wallis' nonparametric H test for numeric variables. The validity of depression diagnoses in the register was determined using the positive predictive value (PPV) defined as the proportion of patients with a diagnosis justified by hospital record review out of the total study sample with a depression diagnosis. Binary logistic regression analysis was used to identify predictors of validity, including the calculation of odds ratios (OR) with 95% confidence intervals (CI) and p-values. All predictors of validity that were statistically significant (p-value <0.05) in the binary logistic regression analyses were included in a multiple logistic regression analysis. McNemar's X^2 test and Cohen's kappa were used to compare the validity of depression diagnoses based on DSM-5 criteria with that of diagnoses based on ICD-10 criteria. All statistical analysis were performed using STATA 16 statistical software (27).

Ethics

The study was approved by the Authority for Patient Safety at the Danish Health and Medicines Authority (journal number 3-3013-2468/1). Data collection and data storage were approved by the Danish Data Protection Agency (journal number 2018-41-5353) and the Region of Southern Denmark's Directory for Processing Personal Data in Relation to Research (case number 18/61816). Furthermore, access to psychiatric hospital record archives was granted by the clinical and research heads of all child and adolescent psychiatric departments.

Results

Sample characteristics

Hospital records were unavailable for N=107 (21.4%) of the 500 randomly selected cases. Unavailable hospital records did not differ from available records regarding age at diagnosis (p=0.617), sex (p=0.981), or depressive episode severity subtype (p=0.625). However, the number of unavailable records differed between the five Danish geographical regions, ranging between 3.3% (Region of Northern Denmark) and 43.3% (Region of Zealand) (p<0.001). This difference was due to some departments using a remote storage facility for hospital records that meant they were not accessible. The year of diagnosis also affected availability, with more of the older hospital records being unavailable (p<0.001).

A total of N=393 hospital records were accessible for the study. The baseline characteristics for these patients (obtained directly from the DPCRR) are reported in Table 1. The study participants from the five geographical regions did not differ as regards sex (p=0.641), age at diagnosis (p=0.683), or depressive episode severity (p=0.211). There were, however, statistically significant differences between geographical regions regarding year of diagnosis with the median year of diagnosis ranging from 2009 to 2013 (p<0.001). There were also differences between geographical regions in terms of the proportions of in-patients, out-patients, and emergency department patients (p<0.001) and whether the depressive episode diagnosis was assigned as a primary or secondary diagnosis (p=0.003). Finally, the number of randomly selected cases varied due to different population sizes of the five regions (p<0.001).

Table 1 approximately here

Validity of depression diagnoses

The raters found that 287 records out of 393 sufficiently documented the presence of depressive symptoms and functional impairment to fulfil the ICD-10 diagnostic criteria for *any* depressive episode diagnosis. This resulted in an overall agreement between raters and register diagnoses of 72.8% (95% CI 68.1-77.0%).

DSM-5 criteria are not used in the DPCRR. Nevertheless, the raters found that 73.3% (95% CI 68.7-77.4%) of the patients assigned with a depressive episode diagnosis also fulfilled the DSM-5 diagnostic criteria for *any* major depressive disorder. Additional analyses showed that there was no statistically significant difference between the validity of the depressive episode (ICD-10) and the major depressive disorder (DSM-5) diagnosis (McNemar's chi^2=0.10, p=0.875).

Cohen's kappa was used to determine whether the two diagnostic classification systems classified the same patients as having depression and not just the same proportion of patients. The agreement was 89.8% and the kappa value was 0.74 (p<0.001), showing a substantial agreement between the two diagnostic systems.

Predictors of validity

The validity was positively correlated with increasing severity of the depression diagnosis as registered in the DPCRR. The validity ranged from 59.3% for mild depressive episode, to 75.8% and 75.9% for moderate and severe depressive episode, to 84.6% for severe depression with psychotic symptoms. Logistic regression analysis confirmed that severity was a predictor of higher validity (p=0.014). The validity of different levels of severity were compared, and moderate depression (OR 2.15, 95% CI 1.27-3.63, p=0.004), severe depression without psychotic symptoms (OR 2.16, 95% CI 1.01-4.62, p=0.046) and with psychotic symptoms (OR 3.77, 95% CI 1.20-11.91, p=0.023) all demonstrated a validity that was higher than that of mild depression. There was no statistically significant difference, when comparing the validity of moderate depression with severe depression (p=0.981), moderate depression with severe depression with psychotic symptoms (p=0.318), and severe depression with or without psychotic symptoms (p=0.377).

Admission status was a statistically significant predictor of validity (Table 2), with the validity of the diagnosis being higher in in-patients than in out-patients and emergency department patients. Diagnoses assigned to out-patients also presented a higher validity than those assigned to emergency department patients.

The validity of depression diagnoses was higher for patients treated with any psychotropic drug than for patients with no psychotropic treatment (Table 2). The diagnostic validity was also higher for patients treated with antidepressants compared to patients treated without medication or patients treated with non-antidepressant psychotropics.

The use of a semi-structured interview or questionnaire in the diagnostic assessment increased diagnostic validity at a statistically significant level (Table 2).

Table 2 approximately here

The validity of the depression diagnosis increased throughout the period under study (Figure 1). Logistic regression analysis confirmed that the year of diagnosis was a statistically significant predictor of validity per year (OR = 1.14, 95% CI 1.08–1.20, p<0.001).

A binary logistic regression analysis showed a time trend in the use of semi-structured interviews and questionnaires, with the use increasing with a more recent year of diagnosis (OR 1.24, 95% CI 1.15-1.33, p=0.000). The correlation between year of diagnosis and use of semi-structured questionnaire was tested using Pearson's correlation and found to be highly correlated (r=0.92, p<0.000).

Figure 1 approximately here

The diagnostic validity in the five geographical regions ranged from 70.2% to 75.9%, showing no statistically significant difference between regions (p=0.949). Furthermore, patient age at the time of diagnosis was not a statistically significant predictor of diagnostic validity (continuous variable p=0.367, categorical variable p=0.846), neither was patient sex (p=0.480) or primary vs. secondary diagnosis (p=0.806).

A multiple logistic regression analysis with validity as the outcome was performed including all variables that were statistically significant predictors of high validity in binary logistic regression analyses (disorder severity, year of diagnosis, admission status, use of semi structured interview or questionnaire, use of antidepressants). Two variables remained statistically significant predictors of a high diagnostic validity in the logistic regression model: a more recent year of diagnosis (OR 1.14, 95% CI 1.07-1.20, p<0.001) and use of antidepressant drugs (OR 2.27, 95% CI 1.35-3.82, p=0.002), whereas emergency department admission was a statistically significant predictor of low diagnostic validity (OR 0.33, 95% CI 0.12-0.93, p=0.036).

Table 3 approximately here

Specific validity of depressive episode *severity subtypes*

Throughout the hospital record review, raters also determined whether the documentation justified the *specific* severity subtype (mild, moderate, severe or severe with psychotic symptoms) that was assigned in the register and not simply *any* depression diagnosis. The raters were able to justify the specific severity subtype in 173 out of 393 patients, providing an overall validity of 44.0% (95% CI 39.2-49.0%). When looking at each severity subtype, the PPVs were mild depression 29.1%, moderate depression 49.8%, severe depression without psychotic symptoms 42.6%, and severe depression with psychotic symptoms 46.2%. Regression analyses confirmed that severity was a statistically significant predictor for the validity of the depressive subtypes (p=0.010). However, when comparing the validity between specific severity subtypes, only the difference between mild depression and moderate depression was statistically significant (OR 2.42, 95% CI 1.42-4.12, p=0.001).

Discussion

This study is the first of its kind to investigate the validity of depression diagnoses assigned to children and adolescents in a national patient research register. The validity of depression diagnoses registered in the Danish Central Psychiatric Research Register (DPCRR) between 1996 and 2016 was good, with expert review of psychiatric hospital records verifying almost 75% of depression diagnoses. Diagnostic validity was highest for the more recent diagnoses, and for depressive episodes that required treatment with antidepressants.

Our study showed an overall validity of 72.8% fulfilling the ICD-10 criteria for any depressive episode diagnosis. This is in line with a Danish study of depression diagnoses assigned to adults in the DPCRR, which found an overall validity of 75.4% (18) and with a study on the validity of early onset schizophrenia in the DPCRR that found a validity of 75.3% (23). However, validity has been reported to be slightly higher for other child and adolescent psychiatric diagnoses in the DPCRR. The validity of early onset obsessive compulsive disorder (OCD) diagnoses was found to be 85% (19), and that of hyperkinetic disorder was 86.8% (20). Another study validating the diagnosis of childhood autism reported a validity of 94% with even 97% meeting the criteria for any autism spectrum disorder (ASD) (22). This leaves the validity of childhood depression diagnosis at the lower end compared to other child and adolescent psychiatric disorders but at the same level as depression in adulthood. We believe that the diversity in validity of registered

diagnoses is partly explained by differences in the presentation and nature of the disorders. Depression is an episodic disorder with the risk of symptoms being disregarded due to its internalizing quality, whereas hyperkinetic disorder and ASD are chronic disorders with more explicit presentation. Further, depressive symptoms may occur in individuals with other psychiatric disorders, which could increase the risk of misclassification compared to e.g. psychotic symptoms and hyperactivity. Finally, depression might progress into disorders like bipolar disorder or schizophrenia (28, 29) and this could contribute to the lower validity compared to other disorders.

The Danish health care system uses ICD-10 criteria for diagnostic classification, and this also applies, therefore, to national patient registers. We found that patients, whose ICD-10 depressive episode diagnosis was confirmed to be valid, also fulfilled the DSM-5 diagnostic criteria for major depressive disorder in 89.8% of the cases. A study validating hyperkinetic disorder diagnoses in children also reported a substantial agreement between the two diagnostic classification systems (20).

There was an almost linear correlation between the depression diagnosis validity and the year of diagnosis. This might be due to an increased focus on proper documentation related to use of systematic assessment tools, which increased over the years. It could also be explained by Danish clinicians gaining more experience in using the ICD-10 criteria that were implemented in 1994. In contrast to the present study, studies examining the validity of hyperkinetic disorder (20) and OCD (19) diagnoses did not find a time trend in diagnostic validity. This difference might be caused by different analytical methods (dichotomous time variables in the hyperkinetic disorder study vs. a continuous time variable in the present study). It could, however, also be due to an increased focus on childhood depressive disorders.

This study found that the validity of depression diagnoses was highest among in-patients when compared to out-patients and patients diagnosed in the emergency department. However, only the lower validity of emergency department diagnoses persisted after adjusting for all predictor variables. So far, our study is the only Danish study investigating the validity of child and adolescent psychiatric diagnoses that include emergency department patients. Only half of the diagnoses of depression assigned in emergency consultations were found to be justified at our specialist review. The limited validity of emergency department depression diagnoses has both clinical and research implications. Even under time constraints and often in highly complex

situations, clinicians should consider carefully whether to assign a first onset depression diagnosis at an emergency department visit. Furthermore, researchers should consider excluding depression diagnoses assigned at the emergency department, when creating a register-based study cohort.

A Danish study examining the validity of childhood schizophrenia diagnoses found a higher validity of inpatient diagnoses compared to out-patients (23). However, this study only performed binary regression analyses, and our results of a higher validity for in-patient diagnoses did not remain statistically significant in the multiple logistic regression model. We believe that this could be due to a correlation between in-patient status and depression severity or use of antidepressants. However, in-patients are observed more closely and by several clinicians (doctors, nurses and other hospital staff), which might lead to more detailed documentation of symptoms.

Our study showed that depression diagnoses assigned for patients treated with anti-depressants had a higher validity than depression diagnoses assigned for patients not receiving anti-depressant treatment. We did not identify other validation studies addressing this important point. It is likely that observations are documented and diagnoses are assigned with extra care, when the decision to initiate medical treatment with possible side-effects is made.

The validity of the depression diagnoses increased with increasing severity of the depression, but the finding did not persist in the multiple logistic regression model. A study on depression diagnoses assigned to Danish adults also observed this tendency in a binary regression model (18). It is fair to assume that more severe cases are more thoroughly examined and present with more explicit symptoms, which could explain this finding. The loss of statistical significance might be explained by more severe cases also receiving more psychotropic treatment.

Our study found that the use of validated assessment tools (semi-structured interviews or questionnaires) increased throughout the study period. We believe that this could have contributed to an improved documentation of depressive symptoms in patient records over time. The use of assessment tools did, however, not in itself increase diagnostic validity in the multiple logistic regression model at a statistically significant level. This was probably due to a strong correlation with recent year of diagnosis, which was a statistically significant predictor of diagnostic validity. A study on the validity of the OCD diagnosis also found an association between the use of a structured assessment tool and increased validity of the diagnosis in a binary regression analysis (19). This highlights the value of these assessment tools in covering the full spectrum of psychopathological symptoms. However, as assessment tools were only used in 27% of the cases

in our study, there is a huge potential for increasing their clinical use and thereby improving diagnostic validity.

There was no geographical variation in diagnostic validity of depression diagnoses between the Danish regions. This finding is similar to that of a Danish validation study of hyperkinetic disorder diagnoses (20), suggesting a nationwide uniformity in the quality of the assigned diagnoses.

Furthermore, we found no association between sex or age at diagnosis with depression diagnosis validity, which is in line with the Danish study of childhood OCD diagnoses (19). This observation suggests that, regardless of age or sex, patients receive the same quality of diagnostic assessment. However, as our sample only included a small number of the youngest patients (32 patients under age 13 years), a correlation with age might have been overlooked.

Finally, our study showed low specific validity for the severity subtypes, with the severity degree only being confirmed in 29% of mild depression diagnoses and 43%-50% of moderate-severe depression diagnoses. This underlines the fact that, even though patients listed as having a depressive episode in the register fulfil the diagnostic criteria for *any* depressive episode, the information about the severity of the depression is less reliable. The validity of specific depression diagnoses is also low for Danish adults, where it ranges between 14-39% (18). The somewhat lower validity in diagnosis of adults compared to that of children and adolescents might be caused by differences in methodology, as the study of adults examined validity by means of patient interview, which could introduce recall bias.

There are limitations to the present study that need to be addressed. The study used information reported in hospital records as a proxy for validity of the depression diagnoses assigned for young psychiatric patients. This procedure implies that symptoms present at the time of diagnosis but not documented will have been missed. However, insufficient documentation would point towards an underestimation of the diagnostic validity reported in our study.

It was not possible to blind raters to the depression diagnosis registered in DPCRR, which could introduce confirmation bias. However, the use of a structured scoring sheet aimed to ensure objective and comparable documentation of the hospital record information available. Furthermore, inter-rater reliability was excellent and assured prior to the study.

We cannot rule out that depression diagnoses registered as ‘first onset episodes’ in the DPCRR were in fact recurrent episodes. However, we believe that the risk for overlooking a previous

depressive episode is minor in this study, because multiple depressive episodes are uncommon before the age of 18 years.

The study had 21.4% missing records, but the attrition analyses did not point to a systematic bias due to this limitation. The validity of depression diagnoses increased with each year of the period under study, However, because more of the older records were missing, the study might have overestimated the overall validity throughout the study period.

A previous study performed post hoc psychiatric interviews to assess diagnostic validity of disorder diagnoses (21). This procedure could, however, raise several methodological issues in our study. First, contacting children and adolescents for interviews regarding a previous depression might be problematic as individuals below the age of 15 years are not qualified to give legal consent in Denmark. Second, contacting adult individuals for interviews regarding a depressive episode occurring in childhood or adolescence could cause significant recall bias. Third, it could cause information bias because the diagnostic procedure in child and adolescent psychiatry relies on a multi-informant approach, collecting information from parents, teachers, and patient. This information would not be possible to collect post hoc. We therefore believe that the methods used in this study are suitable for addressing the research question in this age group.

A major strength of our study is the inclusion of a large representative sample of all children and adolescents diagnosed with a depressive episode in Danish child and adolescent psychiatric departments between 1996 and 2016. Furthermore, trained child and adolescent psychiatrists and the first author carefully reviewed and rated all available hospital records systematically using structured score sheets, which resulted in excellent inter-rater reliability.

In conclusion, this is the first study to examine the validity of diagnoses for depression assigned to children and adolescents referred to psychiatric services. The study found a satisfactory validity of diagnoses of childhood depression in the Danish Psychiatric Central Research Register according to both ICD-10 and DSM-5 diagnostic criteria, confirming a depressive disorder in three out of four cases. Diagnostic validity was highest for the more recent cases and for those receiving treatment with antidepressants. Diagnostic validity was not affected by age, sex, or geographical region.

This study supports the use of data from the DPCRR in future research into depression in children and adolescents. Researchers and clinicians should be aware of the low validity of emergency

department diagnoses, the limited validity of specific subtypes of severity, and of the fact that diagnostic validity of diagnoses for depression has increased linearly over the last two decades.

X. References

1. JAMES SL, ABATE D, ABATE KH, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018;392:1789-858.

2. DALSGAARD S, THORSTEINSSON E, TRABJERG BB, et al. Incidence Rates and Cumulative Incidences of the Full Spectrum of Diagnosed Mental Disorders in Childhood and Adolescence. JAMA psychiatry. 2020 Feb 1;77:155-64.

3. SKOVLUND CW, KESSING LV, MORCH LS, LIDEGAARD O. Increase in depression diagnoses and prescribed antidepressants among young girls. A national cohort study 2000-2013. Nordic journal of psychiatry. 2017 Jul;71:378-85.

4. SCHMIDT M, PEDERSEN L, SORENSEN HT. The Danish Civil Registration System as a tool in epidemiology. Eur J Epidemiol. 2014 Aug;29:541-9.

5. SCHMIDT M, SCHMIDT SA, SANDEGAARD JL, EHRENSTEIN V, PEDERSEN L, SORENSEN HT. The Danish National Patient Registry: a review of content, data quality, and research potential. Clinical epidemiology. 2015;7:449-90.

6. PEDERSEN CB. The Danish Civil Registration System. Scandinavian journal of public health. 2011 Jul;39:22-5.

7. MORS O, PERTO GP, MORTENSEN PB. The Danish Psychiatric Central Research Register. Scandinavian journal of public health. 2011 Jul;39:54-7.

8. MARET-OUDA J, TAO W, WAHLIN K, LAGERGREN J. Nordic registry-based cohort studies: Possibilities and pitfalls when combining Nordic registry data. Scandinavian journal of public health. 2017 Jul;45:14-9.

9. KESSING LV, RYTGAARD HC, GERDS TA, BERK M, EKSTROM CT, ANDERSEN PK. New drug candidates for depression - a nationwide population-based study. Acta psychiatrica Scandinavica. 2019 Jan;139:68-77.

10. WESSELHOEFT R, PEDERSEN CB, MORTENSEN PB, MORS O, BILENBERG N. Gender-age interaction in incidence rates of childhood emotional disorders. Psychol Med. 2015 Mar;45:829-39.

11. RANNING A, BENROS ME, THORUP AAE, et al. Morbidity and Mortality in the Children and Young Adult Offspring of Parents With Schizophrenia or Affective Disorders-A Nationwide Register-Based Cohort Study in 2 Million Individuals. Schizophrenia bulletin. 2019 Jun 8.

12. MUNK-JORGENSEN P, OSTERGAARD SD. Register-based studies of mental disorders. Scandinavian journal of public health. 2011 Jul;39:170-4.

13. WORLD HEALTH ORGANIZATION (WHO). The ICD-10 classification of mental and behavioral disorders: diagnostic criteria for research. Geneva: World Health Organization; 1993.

14. BYRNE N, REGAN C, HOWARD L. Administrative registers in psychiatric research: a systematic review of validity studies. Acta psychiatrica Scandinavica. 2005 Dec;112:409-14.

15. ALLEBECK P. The use of population based registers in psychiatric research. Acta psychiatrica Scandinavica. 2009 Nov;120:386-91.

16. MUNK-JORGENSEN P, OKKELS N, GOLBERG D, RUGGERI M, THORNICROFT G. Fifty years' development and future perspectives of psychiatric register research. Acta psychiatrica Scandinavica. 2014 Aug;130:87-98.

17. PARKER G. Register now: validity later. Acta psychiatrica Scandinavica. 2005 Dec;112:407-8.

18. BOCK C, BUKH JD, VINBERG M, GETHER U, KESSING LV. Validity of the diagnosis of a single depressive episode in a case register. Clinical practice and epidemiology in mental health : CP & EMH. 2009 Feb 12;5:4.

19. NISSEN J, POWELL S, KOCH SV, et al. Diagnostic validity of early-onset obsessive-compulsive disorder in the Danish Psychiatric Central Register: findings from a cohort sample. BMJ open. 2017 Sep 18;7:e017172.

20. MOHR-JENSEN C, VINKEL KOCH S, BRICIET LAURITSEN M, STEINHAUSEN HC. The validity and reliability of the diagnosis of hyperkinetic disorders in the Danish Psychiatric Central Research Registry. Eur Psychiatry. 2016 May;35:16-24.

21. KESSING L. Validity of diagnoses and other clinical register data in patients with affective disorder. Eur Psychiatry. 1998 Dec;13:392-8.

22. LAURITSEN MB, JORGENSEN M, MADSEN KM, et al. Validity of childhood autism in the Danish Psychiatric Central Register: findings from a cohort sample born 1990-1999. Journal of autism and developmental disorders. 2010 Feb;40:139-48.

23. VERNAL DL, STENSTROM AD, STAAL N, et al. Validation study of the early onset schizophrenia diagnosis in the Danish Psychiatric Central Research Register. European child & adolescent psychiatry. 2018 Aug;27:965-75.

24. HARRIS PA, TAYLOR R, THIELKE R, PAYNE J, GONZALEZ N, CONDE JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. Journal of biomedical informatics. 2009 Apr;42:377-81.

25. COHEN J. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement. 1960;20:37-46.

26. LANDIS JR, KOCH GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33:159-74.

27. Stata Statistical Software. Release 15 ed. College Station, TX: StataCorp LLC: StataCorp; 2017.

28. MUSLINER KL, MUNK-OLSEN T, MORS O, ØSTERGAARD SD. Progression from unipolar depression to schizophrenia. Acta psychiatrica Scandinavica. 2017 Jan;135:42-50.

29. MUSLINER KL, ØSTERGAARD SD. Patterns and predictors of conversion to bipolar disorder in 91 587 individuals diagnosed with unipolar depression. Acta psychiatrica Scandinavica. 2018 May;137:422-32.

Table 1 Characteristics of study participants with available hospital records[†]

	Region of Northern Denmark	Central Region of Denmark	Region of Southern Denmark	Region of Zealand	Capital Region of Denmark	Total	P-value
Patients, n (%)	29 (7.3)	86 (21.9)	109 (27.7)	38 (9.7)	131 (33.3)	393 (100)	<0.001
Female sex (%)	62.1	66.3	70.6	68.4	74.1	70.0	0.641
Age at diagnosis, median (range)	15 (11-17)	15 (7-17)	16 (10-17)	16 (8-17)	15 (7-17)	15 (7-17)	0.683
Year of diagnosis, median (range)	2010 (1997-2016)	2009 (1996-2016)	2010 (1997-2016)	2013 (2008-2016)	2011 (2002-2016)	2011 (1996-2016)	<0.001
Severity (%)							0.211
Mild	17.2	18.6	22.9	34.2	20.6	21.9	
Moderate	58.6	59.3	61.5	50.0	55.7	57.8	
Severe	20.7	14.0	13.8	13.2	12.2	13.7	
Severe + psychotic	3.5	8.1	1.8	2.6	11.5	6.6	
Admission status, %							<0.001
In-patients	17.2	10.5	7.3	10.5	19.9	13.2	
Out-patients	82.8	73.3	80.1	89.5	58.8	74.8	
Emergency Department patients	0.0	16.3	4.6	0.0	21.4	12.0	
Depression as primary diagnosis, %	86.2	86.1	89.0	63.1	75.6	81.2	0.003

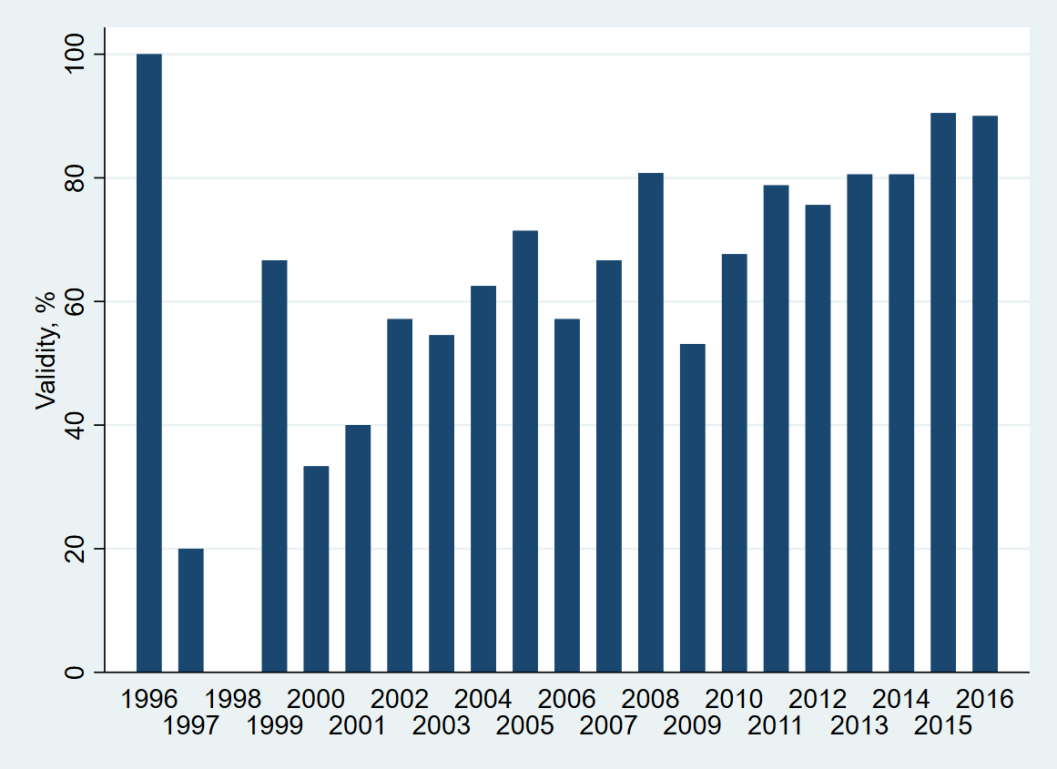
[†] All data was obtained directly from the Danish Central Psychiatric Research Register. Missing hospital records was excluded from statistical analysis.

Table 2 Predictors of validity

Predictor	% of cases	Validity, %	OR	95% CI	p-value
Admission status					0.001
In-patients	13.2	86.5			
Out-patients	74.8	73.5			
Emergency department patients	12.0	53.2			
In-patients vs. out-patients			2.32	1.00-5.36	0.049
In-patients vs. emergency department patients			5.66	2.12-15.09	0.001
Out-patients vs. in-patients			2.44	1.30-4.57	0.005
Treatment with any psychotropic drug	59.6	82.6			
No treatment with psychotropic drugs	40.4	61.2			
Treatment vs. no treatment with psychotropics			2.98	1.88-4.74	<0.001
Treatment with anti-depressants	49.1	81.4			
No treatment with anti-depressants	50.9	64.5			
Treatment vs. no treatment with anti-depressants			2.40	1.51-3.82	<0.001
Use of semi-structured interview or questionnaire	27.0	83.0			
No use of semi-structured interview or questionnaire	73.0	69.0			
Use vs. no use			2.20	1.25-3.87	0.006

Table 3 Results of the Multiple Logistic Regression Analysis

Predictor	OR	95% CI	p-value
Year of diagnosis	1.14	1.07-1.20	<0.001
Treatment with antidepressants	2.27	1.35-3.82	0.002
Severity of depression			
Mild vs. moderate	1.51	0.86-2.68	0.151
Mild vs. severe without. psychotic symptoms	1.71	0.74-3.96	0.209
Mild vs. severe with psychotic symptoms	3.35	0.96-11.60	0.057
Moderate vs. severe without psychotic symptoms	1.13	0.52-2.43	0.760
Moderate vs. severe with psychotic symptoms	2.20	0.66-7.33	0.198
Severe without. psychotic symptoms vs. severe with psychotic symptoms	1.96	0.51-7.51	0.329
Use of semi-structured interview or questionnaire	1.41	0.76-2.64	0.276
Mode of admission			
In-patients vs. out-patients	0.45	0.19-1.09	0.077
In-patients vs. emergency department patients	0.33	0.12-0.93	0.036
Out-patients vs. emergency department patient	0.73	0.36-1.46	0.371



acps_13258_f1.png